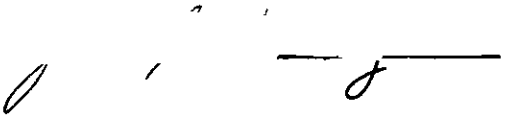
In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.



7/25/68

A MATHEMATICAL MODEL OF AN EPIDEMIC PROCESS

A THESIS

Presented to

The Faculty of the Division of Graduate

Studies and Research

by

Joan S. Horwitz

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in the School of Industrial Engineering

Georgia Institute of Technology

May, 1972

A MATHEMATICAL MODEL

OF AN EPIDEMIC PROCESS

Approved:

Chairman

Date approved by Chairman: May 19, 1972

ACKNOWLEDGMENTS

I gratefully acknowledge the aid, guidance and prolonged assistance of Dr. Douglas C. Montgomery and thank him for introducing and directing me into this area of study. His continual support is indeed appreciated.

I would like to thank Drs. W. W. Hines and H. M. Wadsworth for serving on my reading committee. I am grateful for their suggestions and comments.

My thanks also go to Drs. Timothy A. DeRouen and D. Michael Warner for their many valuable suggestions, criticisms and encouragement.

I wish to express my appreciation and my very special thanks to my husband Stephen for his continued interest, encouragement, and for being so understanding during the time this paper was being prepared, now, and always; and to my parents for their love and encouragement during all of my education.

Finally, I am particularly grateful to Mrs. Helmi Haines for her capable preparation of the illustrations and to Mrs. Shirley Isbell for her expert typing.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Modern medicine is now at a stage where it can do much to relieve or even cure those suffering from infectious diseases. However, its greatest attainments have been in the advancement of preventive medicine, whether these developments be clinical, biological, ecological, or mathematical in nature. It is the mathematical approach, as applied to epidemiology, that will be dealt with here, and it will be specifically applied to the problem of rubella, better known as German measles.

Rubella, which in most cases is of less consequence than the common cold, can sometimes produce serious complications. It is for this reason that the reduction of the number of rubella victims is so important.

Purpose and Scope

The purpose of this research will be to relate both the mathematical theory of epidemics and the epidemiological characteristics of rubella in the form of a mathematical model. Some different approaches will be examined in order to determine which is most conducive to the experimentation desired. These approaches will include both determin-

istic and stochastic models. It is intended that the resulting model be used to describe not only how the population is behaving but also how it will behave in the future and how it should behave under alternative programs of control. In other words, this model is intended to describe, to predict, and to prescribe.

The population under study will be the entire East North Central section of the United States, an area considered very reliable in reporting its rubella cases. There are many drawbacks in trying to fit a mathematical model to a real population, some of which are brought about by the inaccuracies in even the best statistical information. It is for this reason that the initial conditions will be varied to test the sensitivity of the model. Also, alternative vaccination programs will be incorporated into the model to see the effect they have on the continuing rubella process.

Another hindrance is that, as the model includes more factors to more accurately represent the real world phenomenon, the model becomes more difficult to evaluate. This is demonstrated in Chapter VI, which presents an age-stratified model to include various social and biological differences between seven age groups.

Since actual experimentation in the "real world" could not only be time-consuming but also destructive to the population, the "model world" becomes an ideal place in which to experiment.

CHAPTER II

AN INTRODUCTION TO RUBELLA

Rubella was first described in 1815. Termed "roetheln" in Germany (hence the name German measles), the name "rubella" was coined in 1866. The first recorded rubella epidemic was in 1874 (1). Such epidemics were disruptive but did not appear disastrous. Contracted under most conditions, rubella is usually characterized by a few days of a low-grade fever and rash. Although these attacks vary in length and severity, they are seldom cause for much concern. Rubella is highly contagious and is transmitted by nasal secretions. Often, however, infection is not apparent (without symptoms), and transmission usually occurs by direct contact with patients or persons with unapparent infections (2).

In 1941, N. M. Gregg, an Australian opthalmologist, reported the tendency for pregnant rubella victims to give birth to infants with congenital cataracts (1). He also noticed that these infants had an abnormally low birth weight, a high incidence of congenital heart disease, and a high mortality rate. Since that time, Gregg's findings have been confirmed (1). Other effects on the fetus have been discovered. Such harmful results to the child have been termed the "congenital rubella

syndrome" (1). Although the consequences of this syndrome are varied and highly unpredictable, almost every organ of the body may be affected. Congenital rubella syndrome can be characterized by hearing loss, cardiovascular defects, eye defects, developmental and neurological defects, and often cases of spontaneous abortions (1).

It is important to note that a certain percentage of complications other than congenital rubella syndrome can arise. These complications occurring in the infectant are arthritis, encephalitis, and thrombocytopenic purpura (a decrease in blood platelet count) (1).

It was not until 1966 that rubella officially became a nationally reported disease (nationally reported means that each state epidemiologist is required to report the number of cases at the end of each week to the Center for Disease Control (CDC), Atlanta, Georgia). However, due to the lack of side effects associated with the disease, a victim often does not see a physician, especially when it is the second or third infection in a family. In addition, even the medically attended patients are usually not reported to the local or state health department. In fact, it has been estimated by epidemiologists studying rubella that only one case in twenty shows up in the statistics (3).

Prior to 1966, only a small number of states reported their rubella cases (4). Figure 1 depicts the incidence of rubella in some of these states over the past forty years. Keeping in mind the inaccuracy in these data, there still appears to be wide variation in yearly incidence.

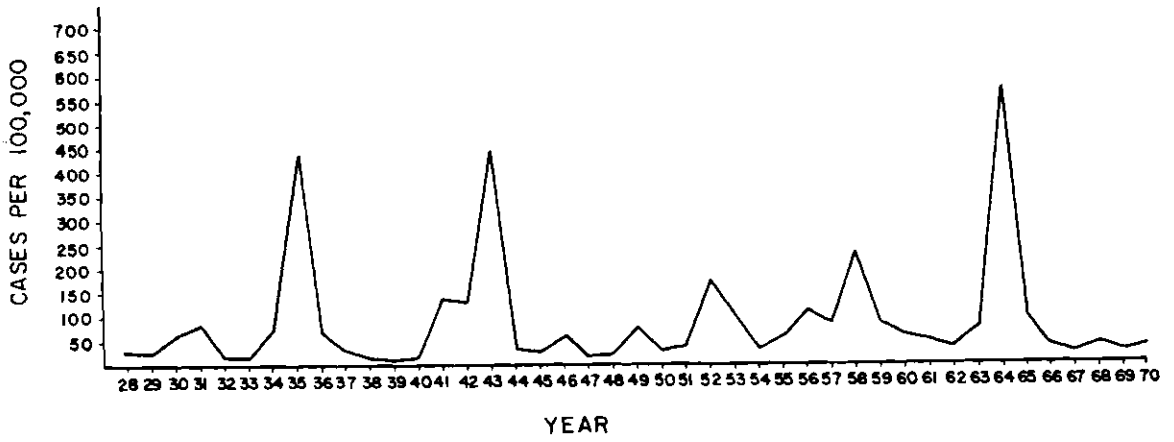


Figure 1. Rubella Incidence (4) - Ten Selected Areas,* U. S. A., 1928-1970.

*Maine, R. I., Conn., N.Y.C., Ohio, Illinois, Wisconsin, Maryland, Washington, Massachusetts

Major epidemics seem to have occurred in 1935, 1943, and 1964, with fairly high incidence rates in 1952 and 1958. These data suggest a period of six to nine years between epidemics.

Figure 2 displays the reported rubella cases by month for 24 selected states (4). Even a cursory glance indicates a strong seasonal pattern within periods. The peaks appear in the spring with the low point in the late summer. It has been suggested that this may be partially due to the dispersion of school children during the summer months.

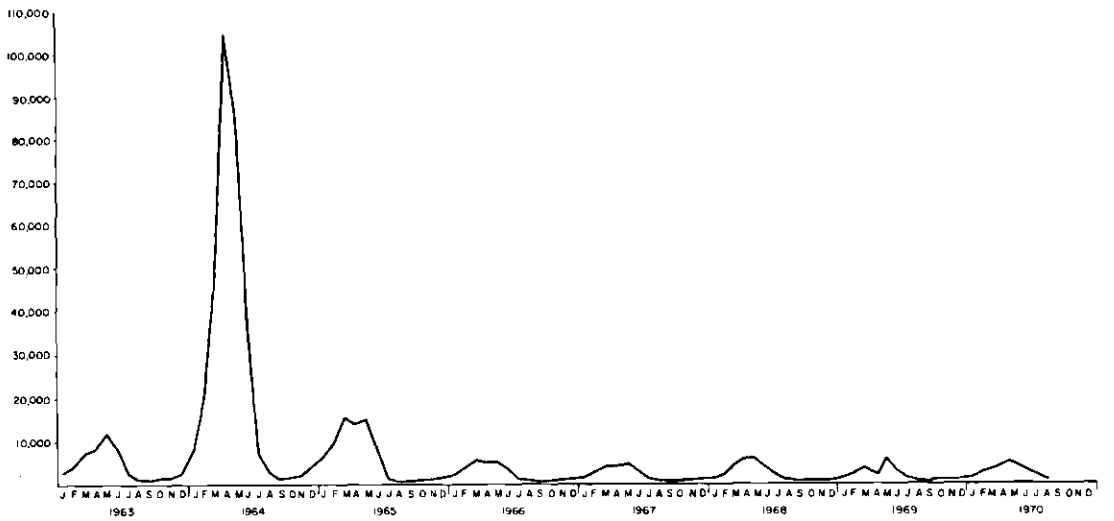


Figure 2. Reported Rubella Cases by Month (4) - Twenty-four Selected States, January, 1963-August, 1970.

Figure 3 depicts the age distribution for reported rubella cases in the years 1963 to 1967, the period of the last major epidemic (4). During this time period, nearly two-thirds of all reported cases occurred in children between the ages of 5 and 14 years. The cumulative distribution indicates that 92 per cent of the cases occurred by age 20. However, there were still a significant number of cases occurring in young adults, particularly women.

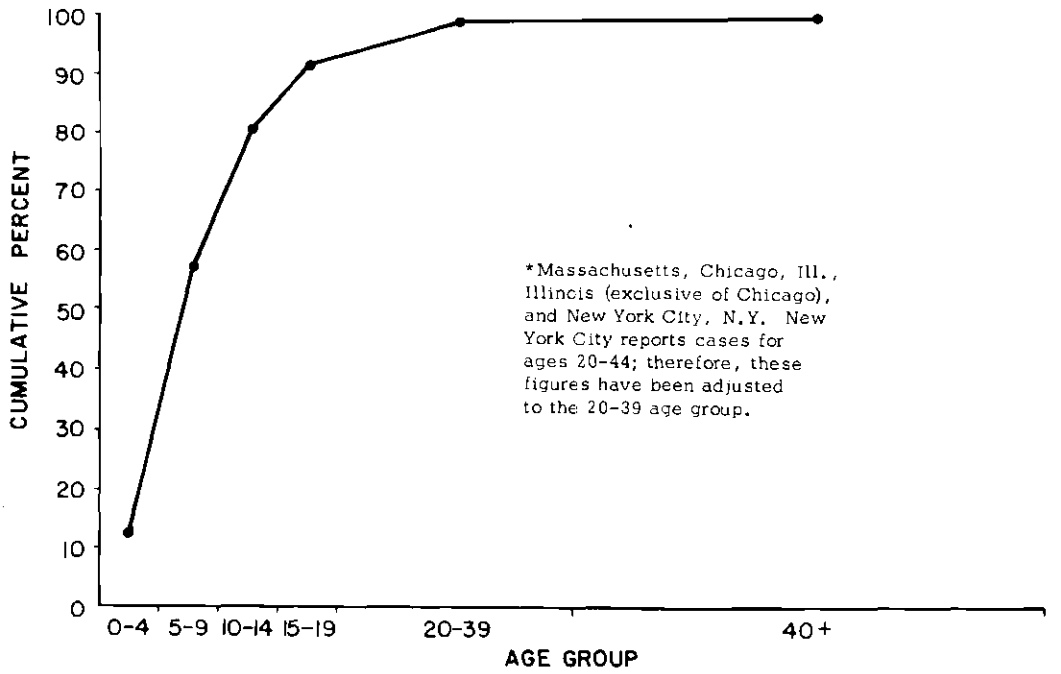


Figure 3. Cumulative Per Cent of Rubella Cases by Age Group (4) (from Selected Areas), 1963-1967.

In the 1964-65 epidemic alone, there were approximately 12,500,000 rubella cases of which 20,000 produced congenital rubella children. Of the 12,500,000 rubella cases, there were 159,375 cases of arthritis, 2,084 cases of encephalitis, 2,160 deaths, 5,000 therapeutic abortions, and 6,250 cases of excess fetal wastage--direct products of the epidemic (4).

It is estimated that this one epidemic cost the American public over $1,461,274,000 in physician and hospital services, loss of potential earnings and special provision for congenital rubella syndrome births in the form of medical and institutional care and special education. In fact, the congenital rubella cases alone cost more than $919,912,000, an average of $45,995 per case (4). It is evident from these figures alone that even a costly means of control might easily be justified.

In 1962, the rubella virus was first isolated. Four years later at the National Institutes of Health, Parkman, et al. (5), were successful in attenuating the rubella virus, enabling them to produce a safe and effective vaccine.

With the introduction of such a successful means of control, it seemed at this point that all problems were solved. At first, however, there was a shortage of vaccine funds. According to an October, 1969, issue of the Journal of the American Medical Association (6), only enough money had been appropriated to vaccinate eight million children

through June, 1970, with a need for fifty to seventy million vaccines to wipe out the infection. Initially, the price of vaccine was high because only one company was licensed to manufacture it. As more licenses were issued, however, the cost was brought down considerably until the cost was no longer an influencing factor.

In 1969, the National Center for Disease Control in Atlanta claimed that a five-year mass vaccine campaign would eliminate rubella. Its strategy was to vaccinate children of at least one year through public school age. It has also been suggested that women of childbearing age also receive the injection to eliminate more quickly the congenital rubella syndrome. However, to avoid any harmful effects to an unborn fetus, this can be done only if no pregnancy can be guaranteed for three months. The American program of vaccinating primarily school children is quite different from that of England. England's policy is to direct vaccination campaigns at the population at risk: women of childbearing age. In fact, in comparison to other diseases the U. S. policy on rubella control is unique in that the target population is not the population most at risk.

Under these constraints it would be helpful to know what results different vaccination campaigns would have on controlling rubella.

CHAPTER III

LITERATURE SURVEY

There is written evidence of epidemic outbreaks of other diseases as far back as the time of the ancient Greeks. However, it was not until the Nineteenth Century that men began to apply mathematical techniques to the statistics of such phenomena.

Curve Fitting

Farr (7), the first to apply such methods, recognized the regularity of the rise and decline of epidemics and that such patterns in epidemic data could be described mathematically. In 1840, Farr used "smoothed" smallpox death frequencies, which he fitted to a normal frequency curve. This was a fairly good representation of the decline of the smallpox epidemic in England, 1837-39. Similarly, in 1866, he took data from the English outbreak of cattle plague and by extrapolating values, he predicted its probable course. Although the curves of his estimations and the actual reported data were similar in shape, Farr's approximation reached a maximum earlier than it actually occurred, with a more rapid decline. His basic assumptions on the orderliness of epidemics were outweighed by the inaccuracy of the curve-fitting technique.

Following in Farr's footsteps, Brownlee (7) fitted Pearson curves to the curves of the epidemics. He found that the epidemic curves were either symmetrical or had a small degree of right skewness. He hypothesized that the lack of negative skewness was due to a decrease in the infecting power of the organism producing the disease. Neither of these curve-fitting approaches, however, has as parameters of the curve the actual factors influencing the epidemic process. With the development of mathematical methods and greater familiarity with epidemics, a new approach was used. Hamer concluded that the epidemic process is dependent upon the number of susceptibles and the rate at which the susceptibles meet and contract the disease from the infectants. It is upon Hamer's classical approach that both deterministic and stochastic models for epidemics are based.

Deterministic Models

In 1911, Ross (8) set up difference equations to model a malaria epidemic. His model included not only the transitions in the human population but also the biological characteristics of the vector (mosquitoes) that spread the disease. He then estimated the rates at which new cases occur and concluded that the number of cases at time $t + 1$ equals the cases at time t plus the new cases minus the recoveries. Although Ross used probabilities to arrive at transfer rates between groups (recovery rate and infection rate), his method remains deter-

ministic in nature.

Similarly to Ross, Kermack and McKendrick (9) used difference equations in their studies. Adding a greater degree of generality, they explored the effects of temporary immunity and specific and non-specific death rates. Their most significant result was the threshold theorem which states that there exists a critical or threshold density value of susceptibles in order for the epidemic to end. The introduction of infectants into such a community of susceptibles would not give rise to an epidemic outbreak as long as the density of susceptibles remained below this critical value. In other words, at equilibrium either an inflow of susceptibles or an increase in the infection rate would tend to favor an epidemic outbreak.

Soper (7), in 1929, refined Hamer's model applying differential equations to measles and setting a time unit equal to the length of the incubation period. Also dealing with a steady state situation, Soper's model was represented by the equation

$$Z = \frac{x}{m} Z_{-1} \tag{3.1}$$

where

Z = cases at time t_o,

Z_{-1} = cases at time t_{-1},

x = susceptibles at time t_o,

m measures "contact rate" and is the steady state number of susceptibles so that one case at time t_{-1} will produce one new case at time t_o.

The difference in time t_o and time t_{-1} is one incubation period. Soper also took into consideration seasonal variations and estimated average monthly values for the contact rate to account for the recurring monthly variation.

Stochastic Models

Up until this point the models were deterministic in nature. In other words, the "future state of the epidemic process can be determined precisely when given the initial number of susceptibles and infectious individuals together with the attack-, recovery-, birth-, and death-rates"(8). The use of stochastic models became more prominent with the realization that there exist random factors capable of altering the course of the disease. The stochastic model deals with the probability or chance of one or more new infectants occurring within a time interval.

Lowell J. Reed and Wade Hamilton Frost (10) modified Soper's model to take into account the fact that contact between one susceptible and one or more infectants will produce exactly one new case. If

p = probability of "sufficient" contact between any two specified susceptibles and infectants,

then

q = 1 - p = probability of this specified contact not occurring.

It follows that if

C_t = the number of infectants at time t,

then

q^{C_t} = probability any given susceptible will not have sufficient contact with any of the infectants

$1 - q^{C_t}$ = probability of sufficient contact with at least one infectant resulting in one infection.

Therefore, if there are S_t susceptibles at time t, then C_{t+1}, the expected number of cases at time $t+1$ is $S_t (1 - q^{C_t})$.

Mathematically this relationship of the probability of C_{t+1} infectants, given S_t susceptibles and C_t infectants at time t, can be described by the binomial probability

$$P(C_{t+1} \mid S_t, C_t) = \binom{S_t}{C_{t+1}} (1 - q^{C_t})^{C_{t+1}} (q^{C_t})^{S_{t+1}} \qquad (3.2)$$

where $S_t = S_{t+1} + C_{t+1}$ since C_{t+1} come from S_t, the remainder of which remain susceptible at time $t+1$. The mean or expected value of this distribution is again $S_t (1 - q^{C_t})$ with variance $S_t (1 - q^{C_t}) q^{C_t}$.

Epidemics can be calculated from this model in both a deterministic and stochastic fashion. Using the deterministic method, C_{t+1}, the number of cases at each time $t+1$, is the mean of the distribution. Incrementing through time

$$C_{t+1} = S_t (1 - q^{C_t})$$

$$S_{t+1} = S_t - C_{t+1}$$

$$S_{t+1} = S_t(1 - 1 + q^{C_t}) = S_t(q^{C_t}).$$

Using a stochastic approach, on the other hand, the outcome is not so uniquely determined. With the aid of a cumulative binomial distribution F(C) and random numbers, the number of cases each time period is generated randomly from a binomial distribution with mean $S_t(1 - q^{C_t})$ and variance $S_t(1 - q^{C_t})q^{C_t}$. Since F(C) is defined over a range 0 to 1, a random number r_o over the same range can be generated so that it is possible to find the value C_o, number of cases, that corresponds to r_o by the inverse function of F.

This stochastic approach to the Reed-Frost model was developed by Helen Abbey and is one of the two forms of the "chain binomial" models. The Helen Abbey model is analogous to the Reed-Frost deterministic model in that it takes into account the fact that contact between one susceptible and one or more infectants can produce only one new case, at the most. Similarly to Hamer-Soper, on the other hand, Greenwood (8) did not make this assumption in his chain binomial model. Greenwood let

> p = the probability of an infectious contact between a case and a susceptible, and
>
> q = 1 - p = the probability of such an infectious contact not occurring.

In this model, therefore, the probability of C_{t+1} cases at time t + 1, given C_t cases and S_t susceptibles at time t is a binomial

$$P\,(C_{t+1} \mid S_t,\ C_t) = \binom{S_t}{C_{t+1}}\, p^{C_{t+1}}\, q^{S_{t+1}} \qquad (3.3)$$

and can be analyzed similarly to the Helen Abbey (Reed-Frost) model.

These models are also called "family" epidemic models because they are usually applied to small groups or to a household. For example, in a household composed of four persons, a double introduction (the first two cases occur at the same time) could produce the following types of chains: (2), (2, 1), (2, 1, 1), or (2, 2). By substituting into both models, the following table of frequencies result:

Total # of Cases	Types of Chain	Probability of Specified # of Cases	
		Greenwood	Reed-Frost
2	(2)	q^2	q^4
3	(2, 1)	$2pq$	$2p(1+q)q^2$
4	(2, 1, 1)	$2p^2q$	$2p^2q^2(1+q)$
	(2, 2)	p^2	$p^2(1+q)^2$

In the Greenwood Model the expected size of the epidemic is 3.78125 and the expected size of the Reed-Frost is 3.96289. In comparison, the Reed-Frost and Greenwood are quite close to each other when computing the total expected size of the epidemic but are not when considered chain by chain (11).

Relying heavily on Kermack and McKendrick's models,

Bartlett (12) developed a simple stochastic model and went on to include the effect of immigration growth processes, prey-predator interactions, and estimation of total epidemic size. In examining deterministic models, Bartlett found that it could be used as an approximation to stochastic models, and as the population size increased, the stochastic results approached the deterministic. Bartlett, in working with recurrent models where the susceptible population is replenished, found that "outbreaks tend to be repeated at intervals having a certain probability distribution" (12).

Bailey (13), in developing a more elaborate model, worked with latent and infectious periods. He found the stochastic approach to be a way of representing a number of epidemics from small groups in a total epidemic curve. Bailey and Bartlett were the source of most of the work in this area during the 1950's.

Recent Results

More recently epidemic models have been developed as prescriptive models. Written after the 1968 epidemic of the Asian flu, Michael Thomas (14) formulated a stochastic model to evaluate the effectiveness of isolation upon the course of the flu epidemic. His model is described by N nonlinear differential difference equations (where N represents the population size) having time-varying coefficients. In order for there to be a closed form solution, N is restricted to rela-

tively small populations such as a college campus or a military training base. Initially, Thomas considered the entire population to be susceptible to the disease and described three distinct phases of the flu, latent, infected, and recovering. Another initial assumption is that the time interval spent in each of these phases is a random variable from a negative exponential distribution. Using a GPSS simulation of the system, Thomas attempted to determine what effect various policies might have in curtailing the epidemic. A surprising conclusion from this analysis was that in the case of the Asian flu, the spread of the epidemic was not deterred by isolating those persons having flu symptoms.

A second prescriptive model was drawn up by Revelle, Feldmann, and Lynn (15). Based on Revelle's doctoral dissertation, this study is an optimization procedure for tuberculosis control. It describes a TB system and then selects types and intensities of controls that will result in a particular number of cases at a least cost.

Initially the population is considered to be divided into nine categories ranging from susceptibles to inactive cases to active cases to recovered active cases. The basic system is complicated by the possible application of chemotherapeutic and prophylactic measures to susceptibles, inactive, and active cases. As in Thomas's study, the model is based on the number of individuals in each category and the transfer of these individuals from one category to another. The system is represented by nine differential equations describing the move-

ment between groups and the birth and death processes. Each time interval is set equal to one year. Revelle, et al. (15), based infection rate along with other parameters on studies done by both Waaler and Ferebee (16) (17). They chose initial conditions to be representative of conditions in some developing nations. Because of the difficulty in solving the difference equations, an approximate method is used, which provides close approximations to the actual solutions.

To represent this as an optimization model, a system of eight linear equations in twelve unknowns was chosen to describe the constraints. The objective function to be minimized was

$$Z = c_m v_m + c_v \sum_i^n v_i + c_g \sum_i^n g_i + c_k \sum_i^n k_i + c_f \sum_i^n f_i$$

where

c_m = cost per vaccination for v_m vaccinations in a mass campaign undertaken in the first year. (Note that v_m must be zero in all years except the first.)

c_v = cost per vaccination for v_i vaccinations in year i.

c_g = cost per vaccination for g_i individuals in the inactive (never having BCG vaccination) group receiving prophylaxis.

c_k = cost per vaccination for k_i individuals in the inactive (having had BCG vaccination) group receiving prophylaxis.

c_f = cost per vaccination for curing f_i active cases.

The objective function represents the total cost for n years. Four plans for reduction of tuberculosis were considered, each spanning a twenty-year horizon period. The four plans differed in the number of years

that the vaccinations were applied and were compared with the efficacy of BCG equal to 0 per cent, 30 per cent, and 70 per cent. When the vaccination program took place in a fewer number of years, the result was a fewer number of man-years of TB in all cases although the costs involved were considerably higher.

Summary

Mathematical epidemic theory has progressed steadily, from fitting known mathematical curves to epidemic curves, to more sophisticated models in which the epidemic process is expressed as a mathematical relationship between the number of susceptibles, infectants, and total population size.

Research has been directed towards both developing more sophisticated epidemic theory and applying these results to specific disease data with the intent of predicting the future course of the epidemic process.

CHAPTER IV

DEVELOPMENT OF A RUBELLA MODEL

Introduction

The first step in developing a model of rubella is to describe the disease process. A simple description is shown below in Figure 4.

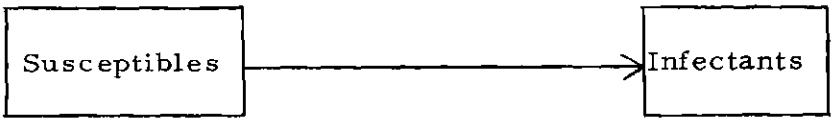


Figure 4. A Simple Epidemic Process.

The population starts out in the susceptible category; in other words, it is not presently infected with the disease but under "proper conditions" is capable of becoming infected. When the individual is exposed to the disease under these "proper conditions," he becomes an infectant. Actually, a distinction should be made at this point between an infectious individual, an infected individual, and an infectant. An infected individual has the disease with symptoms but is not capable of infecting others. On the other hand, an infectious individual is capable of transmitting the disease to others but is not considered to be infected; in other words, he is a carrier. The category infectants combines these two characteristics and is, in this paper, considered synonymous with case. Figure 4,

however, does not describe the complete process. The only flow is from susceptibles to infectants; however, an infectant need not always remain an infectant. There are other alternatives. He may recover from the disease, but then the category he enters depends on the actual process of the disease under consideration. There are several possibilities.

First, upon recovery he may be entirely capable of becoming reinfected. In other words, he returns to the susceptible category until reinfection. This is usually more common among upper respiratory diseases rather than among the "childhood" diseases such as rubella, chicken pox, or mumps. This process is described in Figure 5.

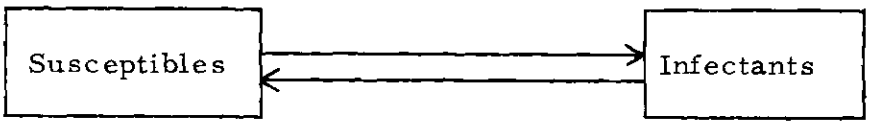


Figure 5. A Simple Epidemic Process with Reinfection.

The other alternative is that the recovered individual becomes immune to the disease and thereby is not capable of reinfection, often known as "superinfection" (11). This process is described in Figure 6.

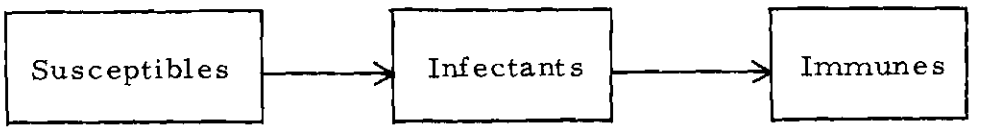


Figure 6. A Simple Epidemic Process with Immunity.

As opposed to Figure 5, here there is no return to the infectant category of individuals who have previously been infectants.

In-between situations may exist either when the recovered individuals have temporary immunity and after a delayed period of time become susceptible, or when the recovered individuals return to being susceptible but are not as prone to reinfection as those who have never been infected. Medically, this might be explained by a higher resistance to infection. This last alternative would be differentiated by a variation in the rate at which the susceptibles without previous infection and the rate at which the susceptibles with previous infection move into the infectant category.

Assumptions and Restrictions as Applied to Rubella

The process shown in Figure 6 is a simple description of the actual rubella process. This system can be described using three differential equations similar in form to those proposed by Kermack and McKendrick (9). They are as follows:

$$\frac{dC}{dt} = \beta C S / P \tag{4.1}$$

$$\frac{dS}{dt} = -(\beta C S / P) - C \tag{4.2}$$

$$\frac{dI}{dt} = C \tag{4.3}$$

where

S = the number of individuals in the susceptible category.

C = the number of individuals in the infectant (or case) category.

I = the number of individuals in the immune category.

P = the number of individuals in the total population ($P = S + C + I$).

β represents a contact rate.

t = time.

$\beta C S / P$ = the rate of infection of susceptibles; it is an output from the susceptibles and an input to the infectants.

Equations 4.1, 4.2 and 4.3 describe an oversimplified system. There are other effects on the rubella process that need to be represented if the model is to be a true picture of the epidemic. These effects include births, deaths, chemotherapy, prophylaxis, immunity, immigration, and emigration.

First, there are both births and deaths taking place in the popu-

lation that are constantly altering the population size. In the case of rubella, all newborns become susceptible (an infant actually retains his mother's immunity for approximately six months; however, for the purpose of the model, this period of temporary immunity will not be included). Therefore, all births will be added to the susceptible population. On the other hand, deaths are from all groups, or the total population. But, whether these deaths come out of each category at the same rate must be considered. In other diseases they may not--the death rate among infectants may be higher than among the well population. Although the chance does exist to die from rubella (as with measles encephalitis), the probability is very small and for the purpose of this model will be considered to be zero. In other words, it will be assumed that rubella does not increase the risk of death and the death rate will be constant among all groups.

A third consideration is chemotherapy, the treatment of disease by chemical reagents. Drug therapy does not alter the course of rubella in terms of a more rapid recovery and is only for the purpose of treating symptoms. Since it does not modify the length of disease nor the infectivity of the disease, chemotherapy will not be a parameter in this model.

On the other hand, prophylaxis does exist in the form of a relatively new rubella vaccine. Although antibody levels (a measure indicating immunity) are generally lower in vaccinated individuals as compared

to those observed in response to natural rubella infection, this vaccine does protect against illness on natural exposure, or adequate contact. According to the Public Health Service Advisory Committee on Immunization Practices (18), such antibody levels in vaccinated individuals have declined slightly over a five-year observation period (although immunity persists), but only continued observation can absolutely determine whether long-term protection can be expected. Rubella vaccine as a form of prophylaxis will be included in this model from 1969 on, since it was not until this time that it came into active use. For the years prior to and including 1968, vaccine will be considered nonexistent. The vaccine and natural infection will be assumed to provide long-term protection, or immunity, against rubella.

As a last consideration, immigration and emigration will not be included because the exact nature of these would be impossible to define. Therefore, the model will be restricted to a closed population.

In review, the assumption and parameters incorporated into this model are as follows:

1. Births are an input to the susceptible population only.
2. Deaths are an output from all categories of the population at the same rate.
3. Chemotherapy will not be included.
4. Vaccine will be included beginning in 1969.

5. Both rubella vaccine and natural infection confer long-term immunity.

6. The model is restricted to a closed population.

7. Homogeneous mixing.

The model can now be described by the following differential equations:

$$\frac{dC}{dt} = \beta C S / P - \gamma C \tag{4.4}$$

$$\frac{dS}{dt} = - \beta C S / P + \mu P - \gamma S - \nu S \tag{4.5}$$

$$\frac{dI}{dt} = C + \nu S - \gamma I \tag{4.6}$$

$$\frac{dP}{dt} = \mu P - \gamma P \tag{4.7}$$

where

S = the number of individuals in the susceptible category.

C = the number of individuals in the infectant (or case) category.

I = the number of individuals in the immune category.

P = the number of individuals in the total population ($P = C + S + I$).

β represents a contact rate between susceptibles and infectants leading to infection of the susceptible individual.

μ = the birth rate.

γ = the death rate.

ν = the percentage of susceptibles being vaccinated.

Method of Solution

Several avenues towards solution of the model described are possible. Often the simplest method of evaluating a set of equations is to solve them mathematically. However, Equations 4.4 through 4.7 do not lend themselves to a direct mathematical solution. In trying to solve similar equations, Bailey (8) found that he had to resort to approximations. The purpose, however, of this research was not to derive a single numerical solution for any of the variables but rather to observe happenings in a model population. It is for this reason that a simulation technique was chosen.

At first a Monte Carlo simulation model was considered. However, this would involve deciding on distributions for each of the random variables (contact rate) and then choosing a random number each time period to determine the value of the random variable. This process is time consuming (on the computer) when at the same time a deterministic approach would be an approximation to such a method.

An example of such a stochastic model would be a variation of the Reed-Frost model (to apply to rubella). It was found, however, that this approach placed limitations on the population size, a condition

not desired in this study. A similar approach was used by Elveback (19) and overall running time was fifty minutes for 100 trial epidemics studying a population of only 500 persons. On the other hand, this method would show the variations in the possible outcomes rather than only mean values.

A computer simulation of a deterministic model was chosen so that large populations could be dealt with and single numerical solutions would not be the only product of the model. Chapter V discusses the source of data or input to the model and the initial status of the population. The model is then simulated using a set of initial conditions, and the results are compared to the observed occurrences in the real population under study. The initial conditions are later varied to test the sensitivity of the model to variations in these values. Also, alternative vaccination programs are incorporated into the model to see the effect they have on the continuing of the rubella process.

Chapter VI deals with an age-stratified epidemic model in which homogeneous mixing among individuals in the population is not a necessary condition.

CHAPTER V

RESULTS OF SIMULATION

The method for solving the model developed in Chapter IV and represented by Equations 4.4, 4.5, 4.6, and 4.7 is a FORTRAN IV simulation of a real population. Written originally for a Burroughs 5500, this program was first modified for an IBM 360, and final runs were made using an IBM 7094. It is displayed in Appendix I.

Source of Data and Initial Conditions

The basis for most of the rubella data used in this study is the Center for Disease Control (CDC), Atlanta, Georgia. In 1966, the Conference of State and Territorial Epidemiologists placed rubella on the list of notifiable diseases. This means that each state is required to report the number of infected cases at the end of each week. Prior to 1966, many states voluntarily reported rubella cases, but this did not necessarily mean that the entire state was reporting. Although data are now submitted in a weekly telegraphic report, the completeness in reporting as well as the type and accuracy of the information varies both between and within states. Based on conversations with rubella epidemiologists and the record librarian at the CDC (who is responsible for this data collection), it is felt that data from certain sections of the

country are more reliable than others--particularly reliable is the East North Central section of the United States.

On this basis the East North Central section was chosen as the test population for study. This section includes: Ohio, Indiana, Illinois, Michigan, and Wisconsin. Due to data availability, the simulation study begins in 1966. Based upon fitting this model to 1966 data, expected data for future years will be predicted.

The model is divided into intervals of two-week periods, the length of time a patient is infectious. Since rubella is characterized by a fairly constant period of infection, discrete time periods is a valid assumption. Another reason for two-week intervals involves reporting. Since rubella is reported weekly, this is a way of smoothing the data. Thus, there will be 26 such intervals per year.

A deterministic approach was chosen since a large population is being considered. Thus, the model is dependent upon the initial conditions and the rates involved.

These initial conditions are the numbers of persons in each of the categories--infectants, immunes and susceptibles--the sum of which is the total population. The total study population is equal to the total population of the East North Central section of the United States at the beginning of the simulated time period, 1966. According to U. S. census data (20), the population at that time was 38,480,000.

The initial number of infectants was based on the number of

reported cases during the first two weeks of 1966. For the East North Central section (ENC), this number was 416 infectants, 103 in Week 1 and 313 in Week 2.

The initial number of immunes, on the other hand, is much more difficult to define. It is not a statistic that is recorded. It cannot be counted on the basis of a person's remembering having a history of the disease. Several stratified random serosurveys (4) have been conducted in different areas of the country to define age-specific immunity on the basis of persons with rubella hemagglutination-inhibition (HI) antibody. Most of these studies have placed the per cent of seroimmunes in the total population, in 1966, somewhere between 75 and 80 per cent. It is important to remember that since rubella reporting is "incomplete and diagnostic accuracy variable, and since a significant proportion of rubella infections are subclinical" (4), serological data is about the only way to determine this information regarding susceptibility. On this basis, the initial number of immunes will be assumed to be 30,000,000 (approximately 78 per cent). There will be further discussion in later sections of this chapter on the sensitivity of the model to variations in this value.

The remaining initial population value is the number of susceptibles. Since $S = P - C - I$, the initial number of susceptibles is 8,479,584 (approximately 22 per cent).

The initial rates to be included are μ, γ, and β. The

birth rate, μ , and the death rate, γ , were based on 1970 census data (20) which give yearly birth and death rates for different sections of the country. Since there was little variation in these rates, let $\mu_y = 19.1$ per thousand where μ_y is the yearly birth rate, and let $\gamma_y = 9.7$ per thousand where γ_y is the yearly death rate. However, since this simulation is based on 26 two-week periods, the birth and death rates need to be biweekly rates. In computing μ and γ , it is important to realize that this has similar properties to a compounding function. In the case of births, the number of people in the population at the end of the first year (not accounting for deaths) is

$$P_1 = (1 + \mu_y) P_o \tag{5.1}$$

where

P_o = the initial population size.

μ_y = the yearly birth rate.

In converting from yearly increments to biweekly increments, the birth process can be represented by the following equation:

$$P_1 = (1 + \mu_y) P_o = (1 + \mu)^T P_o \tag{5.2}$$

where

T = the number of discrete time intervals per year (in this case $T = 26$).

μ = the birth rate per two-week period.

Solving for μ,

$$\mu = e^{(\log (1+\mu_y))\ /\ T} - 1 \qquad (5.3)$$

Similarly, the biweekly death rate can be computed

$$\gamma = 1 - e^{(\log (1-\gamma_y))\ /\ T} \qquad (5.4)$$

where

γ_y = the yearly death rate.

γ = the death rate per two-week period.

In choosing β, the contact rate, it is important to remember that it is a measurement of the susceptibility of the host, the infectivity of the rubella virus, and the social conditions in the population at risk. It is obvious that these are subject to change over time and may be influenced by such factors as children being in and out of school. Therefore, if the reported rubella cases are used to calculate this contact rate, it seems that β varies over time within each yearly cycle. From the epidemiology of rubella, it is known that there is a strong seasonal pattern in reported rubella cases per year. Based on this cyclic variation, β must be time dependent and will be represented by β_t, $t = 1, \ldots T$ where $T = 26$, the number of time intervals considered in one year.

The β_t's will therefore be calculated from 1966 rubella data and then substituted into the model to predict the following years.

The basic equation which determines the number of infectants at each time period is

$$C_{t+1} = \beta_t S_t C_t / P_t - \gamma C_t \qquad (5.5)$$

where

t = a time period t.

Solving for β_t,

$$\beta_t = (P_t / S_t) ((C_{t+1} / C_t) + \gamma) \qquad (5.6)$$

Substituting in initial values for susceptibles, immunes, total population size, death-rates, and the reported number of infected cases per week for one year, all the needed information is known. Thus, the β_t's can be computed for t = 1,...26.

The Initial Model

When simulating the model using these conditions, it was found that the epidemic process did not continue very far into the first year of predictions (1967). In trying to find a cause for the failure of the model, the possibility of under reporting was reconsidered. According to Witte, et al. (3),

> It is estimated that only one case of clinical rubella in 20 is reported. Although these surveillance data cannot be considered quantitatively accurate, it is useful to depict the trends and patterns of rubella occurrence in the United States.

If this is the case, it is reasonable that a model based on such under reporting should quickly diminish since such a small number of cases would not keep the epidemic going.

To test this hypothesis, a correction factor was introduced increasing the reported cases by a multiplicative constant.

The results of a simulation with a correction factor of 22 are displayed in Figure 7 and Table 1. The cases have been accumulated also by 4-week periods in Figure 8 and by 8-week periods in Figure 9 in order to examine what is occurring by months and by seasons. The total number of infectants are displayed as reported. Although the multiplicative constant, FACTOR, has been incorporated into the model, the observed data has been redivided by this constant before being recorded in Figures 7, 8, and 9 and Table 1.

The expected cases (those reported to CDC) in 1967 and 1968 were 8,311 and 11,232, respectively. The observed cases, produced by the model, for these two years were 8,132 and 13,922. It is interesting to note from these graphs that the observed epidemic peaks later than expected, a characteristic also found in Abbey's (10)Reed-Frost application to German measles. In both 1967 and 1968, the rise in observed cases occurs at a slower pace than expected, but upon reaching

TABLE 1. Simulation: Results of Rubella Model

RATE	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5	YEAR 6	YEAR 7
6.071	416	165	187	533	1131	458	172
5.021	555	228	268	786	1605	612	237
6.769	614	259	319	958	1881	676	269
5.517	917	400	513	1577	2966	1008	413
4.921	1116	502	672	2111	3793	1225	518
4.444	1211	563	786	2515	4298	1327	580
4.131	1186	571	831	2696	4365	1297	587
4.886	1079	539	816	2676	4087	1178	552
4.672	1161	602	949	3130	4492	1265	616
4.861	1194	644	1055	3482	4681	1298	657
4.282	1277	718	1221	4008	5031	1385	730
1.506	1202	705	1243	4035	4714	1301	715
4.156	397	243	445	1418	1539	429	246
2.717	363	232	440	1374	1385	391	235
4.902	217	145	285	870	814	234	146
3.140	235	164	333	994	864	252	165
4.620	163	119	250	727	587	175	120
4.232	167	127	276	784	589	178	128
4.661	157	125	281	774	541	167	125
3.634	163	136	314	842	548	173	136
4.069	132	115	275	715	433	140	115
5.530	120	109	270	680	384	127	109
4.897	149	142	360	880	464	157	141
5.062	164	163	427	1008	497	173	161
4.504	187	194	523	1194	550	196	192
3.845	190	206	571	1258	543	199	203
T O T A L	14757	8132	13926	42056	52813	16040	8281

BIRTH RATE= 0.000728 DEATH RATE= 0.000375 FACTOR= 22

IMMUNES	30000000	30031388	29917831	29932209	30561472	31418904	31464561
POPULATION	38480000	38834847	39192980	39554415	39919186	40287319	40658847

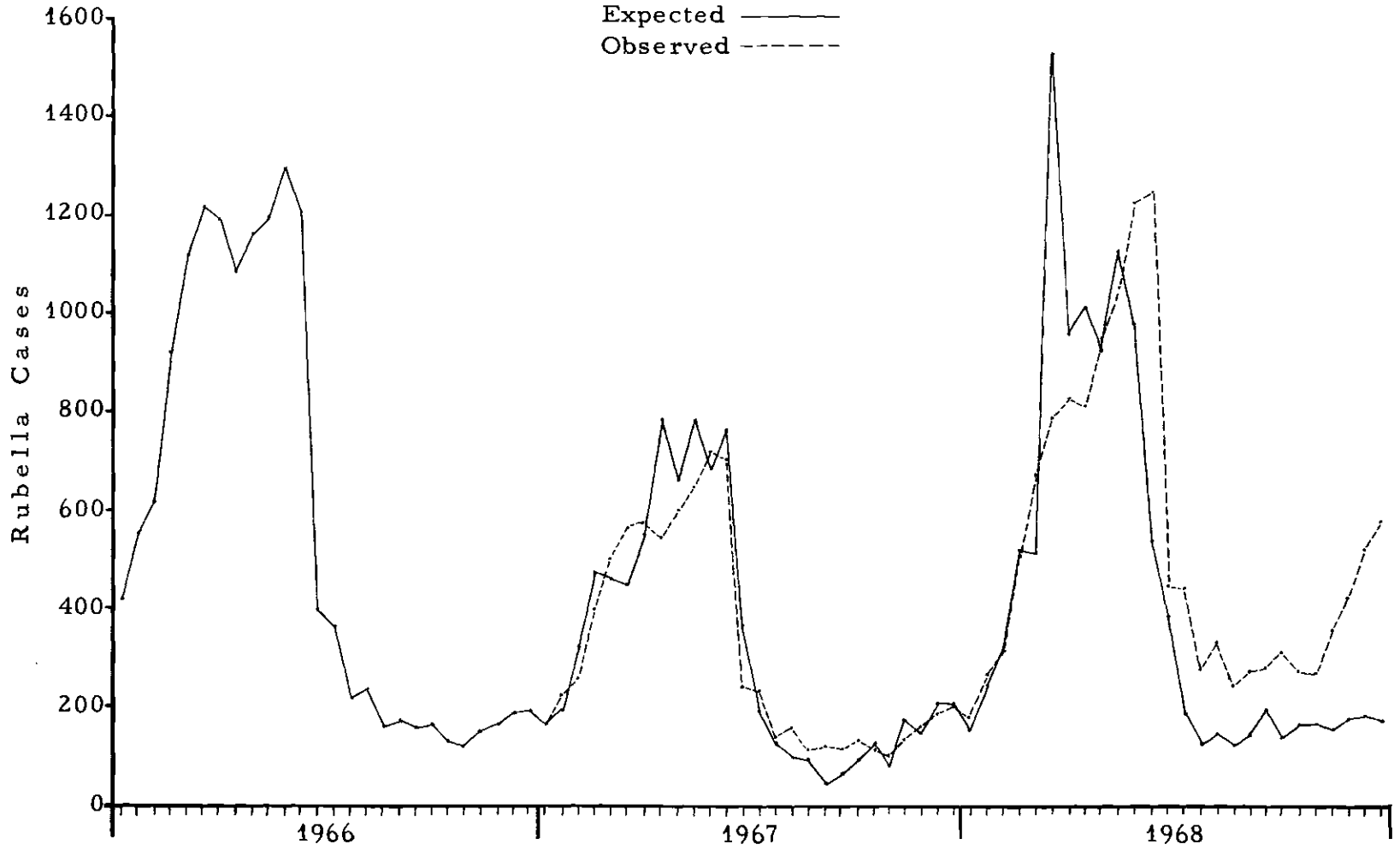


Figure 7. Rubella Data by Two-Week Periods

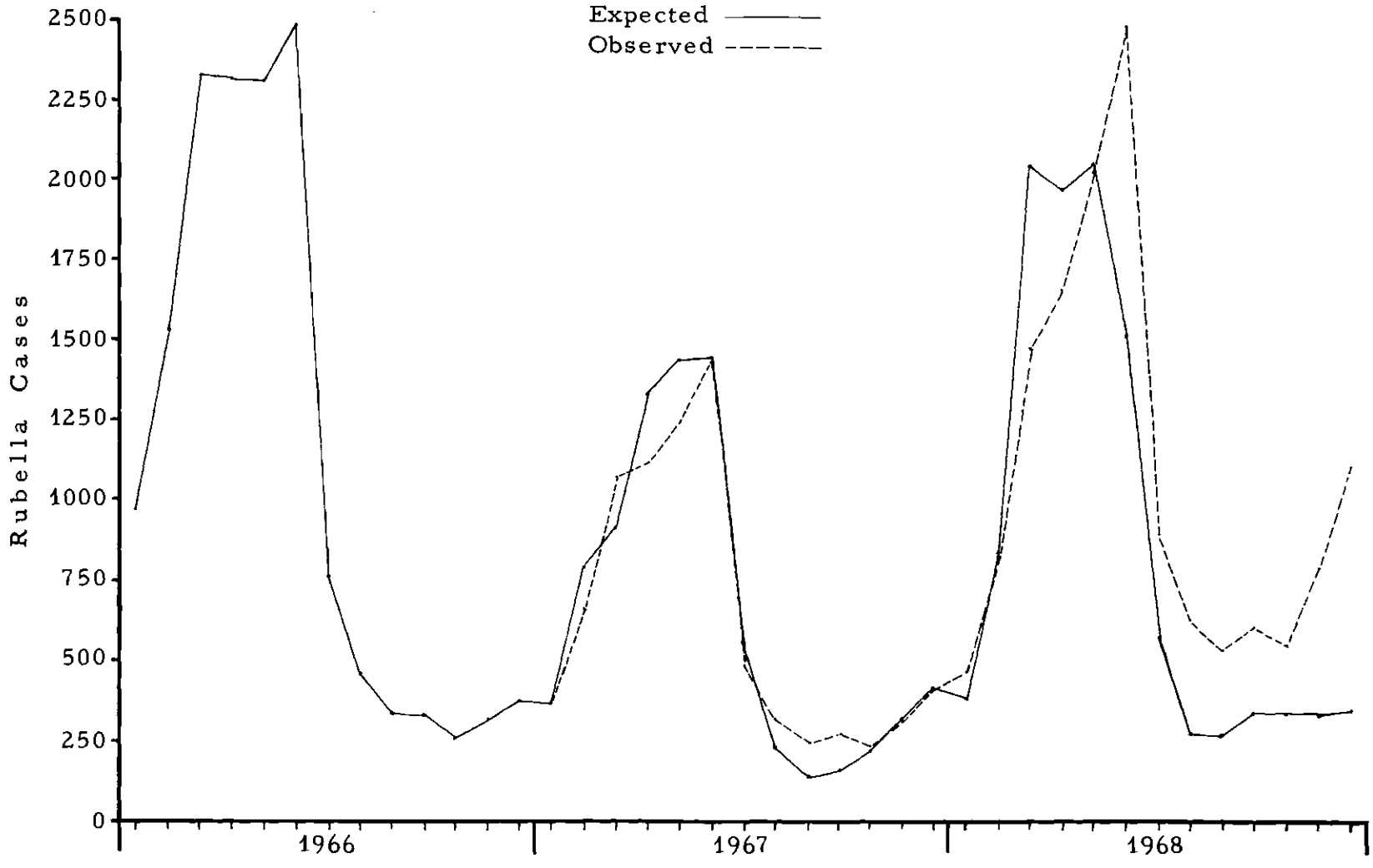


Figure 8. Rubella Data by Four-Week Periods

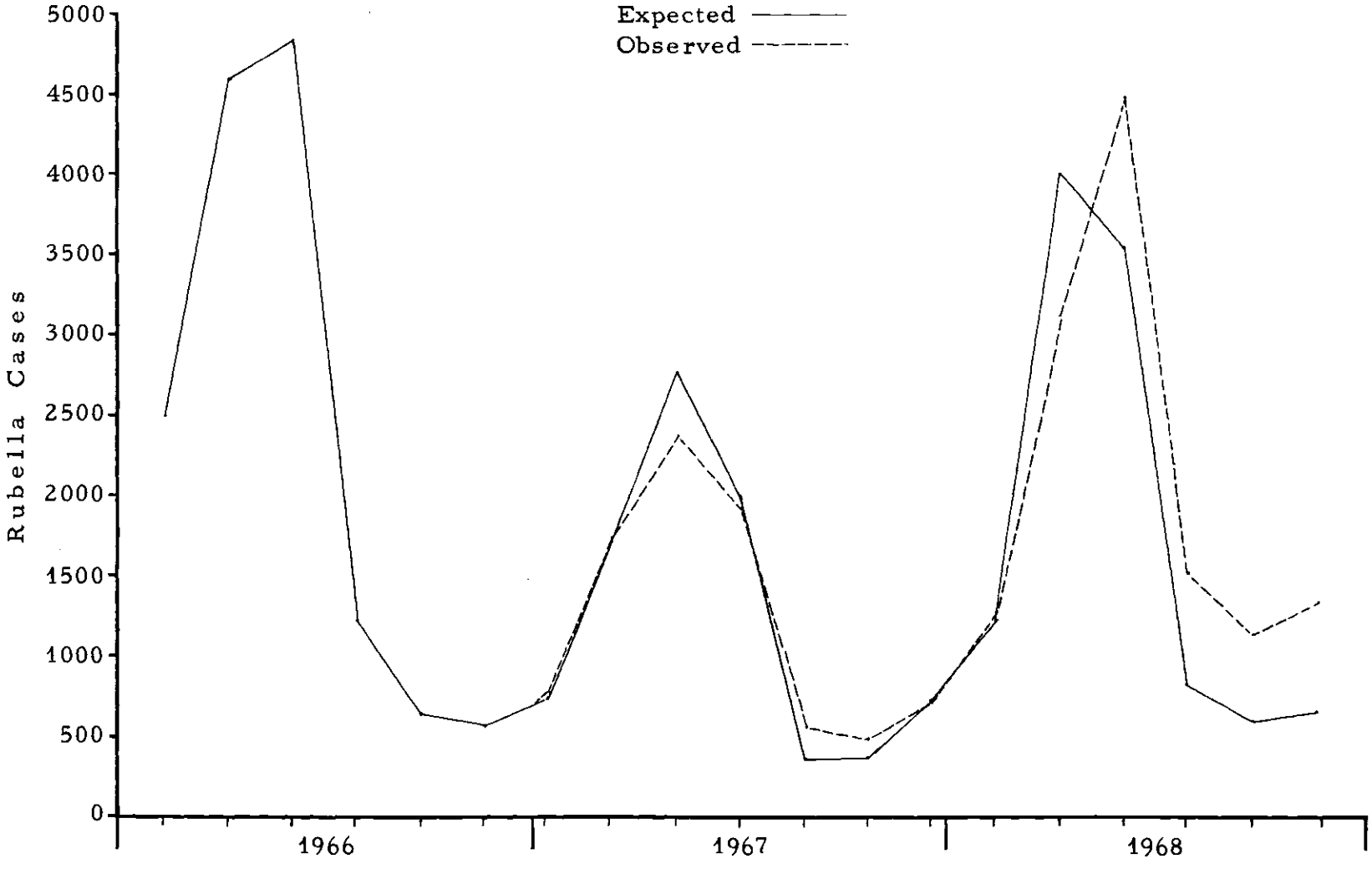


Figure 9. Rubella Data by Eight-Week Periods

a peak, the observed cases decline just as rapidly (and at basically the same point in time) as the expected cases.

The largest discrepancy in 1968 predictions occurs at the latter part of the year, the last four to six months primarily. This might be explained by the fact that vaccination programs (although small in scope) were beginning to occur at this time. Through June 30, 1970, 3,125,925 doses of rubella vaccine were administered through the state and local health departments (21).

The problem with incorporating this knowledge into the model is twofold. First, there is not accurate data regarding the susceptibility of the people receiving the vaccine. For example, the procedure for administering in the schools is to send a note home to the parents requesting permission to vaccinate their children. If the note is signed, the child receives the vaccine. Whether or not the child has had a history of rubella may or may not be known to the parent, and could even have no bearing on the parent's granting permission. In other words, the population receiving prophylaxis may not all be susceptibles. In fact, CDC estimates that 30 per cent of the children receiving immunization are already immune (21).

Second, these three million reflect only vaccinations given through the health departments. Accurate accounting of vaccine given out through private physicians and hospitals is not available.

Sensitivity to Variations in the Initial Number of Immunes

The actual numbers of immunes and susceptibles in a population are difficult values to determine. The problems involved in counting infectants multiply when trying to count immunes as persons remembering having rubella. The only other alternative is to base such estimates on serological surveys, a method quite costly even on a small population.

It would therefore be interesting to see how sensitive the model presented is to variations in the initial numbers of susceptibles and immunes. The simulation discussed assumed 30,000,000 persons (78 per cent) to be immune to rubella. Table 2 shows the results of simulation of the same epidemic process with the initial numbers of immunes equal to 22,000,000 (57 per cent), 25,000,000 (65 per cent) 28,000,000 (73 per cent), and 32,000,000 (83 per cent) again using a multiplicative factor of 22 to correct for under reporting.

It can be seen from Table 2 that as the initial number of immunes increases, the number of cases also increases. At a quick glance, this seems contrary to Equation 5.5, for as the number of immunes increases, the number of susceptibles decreases and the number of infectants should increase respectively. However, the decreasing of susceptibles has actually created larger β_t's initially, an increase sufficient to create more infectants in later time periods. In other words, the increased β_t has caused the product

$$\beta_t S_t$$

to be larger when there are fewer susceptibles.

TABLE 2

A Comparison of Predicted Infectants (Reported) Produced by Varying the Initial Number of Immunes

Initial Number of Immunes	Years 1967	1968
22,000,000	6,707	4,223
25,000,000	6,771	5,682
28,000,000	7,301	8,927
30,000,000	8,140	13,932
32,000,000	9,512	25,486
Actual Data	8,311	11,232

Another interesting observation is that as the initial number of immunes gets larger, the next major epidemic occurs sooner. These results are displayed in Table 3.

Table 3 shows that when there are 22 million immunes (16,479,000 susceptibles) in 1966, the epidemic does not peak until years 1972 and 1973 with an average of 45,526 (times 22) cases per year. When there are 32 million immunes (6,479,000 susceptibles) in 1966, the epidemic peaks as yearly as 1968 and 1969 with an average of 43,512 (times 22) cases in those years.

TABLE 3

Predicted Year of Next Epidemic and Size of the Epidemic Produced by Varying the Initial Number of Immunes

Initial Number of Immunes	Year(s) of Next Epidemic	Average Size of Epidemic
22,000,000	1972, 1973	45,526
25,000,000	1971, 1972	46,602
28,000,000	1970, 1971	46,108
30,000,000	1969, 1970	47,387
32,000,000	1968, 1969	43,512

It might be expected that the trials with more susceptibles should produce an epidemic sooner because it does not need to wait for a build-up of the susceptible population. However, the number of infectants produced under these conditions was not sufficient to spread the disease, and it was actually the infectant population that needed time to accumulate.

Both Tables 2 and 3 demonstrate observations produced through simulation that are not immediately evident from analysis of the model equations. This points up the additional information that can be produced through simulation that cannot be inferred from analysis of descriptive equations alone. Although the model does appear to be sensitive to variations in the initial number of immunes (and susceptibles), this is not necessarily a detriment. Rather it illustrates the importance of fairly accurate information about the population under study.

Sensitivity to Variations in Under Reporting

In addition to the size of the susceptible and immune populations being difficult to determine, the infectant population is not clear cut. Indications of the size of the infectant population at any time period come from the number of reported rubella cases. As was discussed earlier, it is generally agreed that these reported cases are actually only a small percentage of the true number of infectants but are indicative of "trends and patterns of rubella occurrence" (3).

To compensate for this under reporting, a correction factor was introduced. Table 1 showed the results of simulation using a correction factor of 22. (This assumes that only 1 in every 22 cases is actually reported.) Since this was only an hypothesis, other correction factors have also been considered to test the effect these assumptions have on the model. The results of these simulations are shown in Table 4.

It can be seen from Table 4 that although the number of predicted reported cases decreases as the correction factor increases, the predicted cases (the product of predicted reported cases and the correction factor) actually are increasing. However, it appears from Table 4 that these variations do not make the widespread changes in the predicted number of cases as did comparable variations in the number of immunes.

TABLE 4

A Comparison of Predicted Reported Cases of Rubella Produced by Varying the Under Reporting Correction Factor

Years	Under Reporting Correction Factor 15	18	20	22	Actual
1967	8,861 (132,915)	8,540 (153,720)	8,336 (166,720)	8,140 (179,080)	8,311
1968	18,855 (282,825)	16,483 (296,694)	15,134 (302,680)	13,932 (306,504)	11,232

Extent and Timing of Vaccination Programs

The advent of rubella vaccine in the late 1960's has provided a solution to the rubella control problem. At first it seemed that this program was to be stifled purely due to lack of funds. However, by late 1970 enough money was available to complete a mass immunization program consisting of fifty to seventy million vaccines.

Although the United States was able to produce such funds, it is not always as easy for developing nations to do so. With limited resources, it would be advantageous to have a method for testing the effects of alternative control programs. The simulation model presented in this thesis provides the means for such a test.

It will be assumed for the purpose of this model that all those receiving the vaccine are susceptibles. Although this is not usually the case, in cost analysis an adjustment can be made to cover the num-

ber of immunes expected to be vaccinated. It has been estimated (21) that 30 per cent of those receiving vaccine in the East North Central section of the United States through Public Health programs were already immune.

As characterized by Equations 4.5 and 4.6, vaccinated susceptibles are removed from the susceptible population and transferred to the immune population. Since all indications are towards long-term immunity, vaccinated individuals will remain immune.

The control programs under consideration include no vaccination (results displayed in Figure 7), 10 per cent, 25 per cent, 50 per cent, and 75 per cent of the susceptible population. Table 5 summarizes the results of these different vaccination programs, all taking place entirely within year 4, or 1969.

TABLE 5

Rubella Cases Produced Under Alternative Vaccination Programs

Year	Per Cent Susceptibles Vaccinated				
	0%	10%	25%	50%	75%
1969	42,015	32,530	24,177	16,781	12,592
1970	52,760	10,493	516	0	0
1971	16,079	4,460	0	0	0
1972	8,328	7,997	0	0	0
1973	13,628	37,167	0	0	0
Total Size	132,800	92,647	24,693	16,781	12,592

As can be seen from Table 5, it is only necessary to vaccinate 25 per cent of the susceptible population in order to stop the spread of rubella infection in the entire population. It is important to remember, however, that in order to be sure of vaccinating 25 per cent of the susceptible population, it is usually necessary to vaccinate many more individuals because vaccination programs are usually conducted without being sure of each individual's status. Also, infection may reoccur if it is introduced from outside the population under study. Recall that a necessary assumption of this model was that the population be considered closed--no immigration or emigration. If the infection dies out in a closed population, it cannot start again.

It can also be seen in Table 5 that a twofold increase in the per cent vaccinated (in the case of increasing from 25 to 50 per cent) produces only a one-third decrease in the number of cases in the epidemic. It is necessary to triple the 25 per cent program to 75 per cent in order to halve the number of cases, from 24,693 to 12,592.

Timing also plays an important part in the effect vaccination programs have on the course of rubella. Considering a total vaccination program of 25 per cent, Table 6 shows that if the 25 per cent program is delayed just one year, the number of total cases increases from 24,693 to 77,525 and increases to 105,393 if the same program is delayed two years.

TABLE 6

Timing Study with 25 Per Cent Susceptibles Vaccinated

Year	Year Vaccination Takes Place		
	1969	1970	1971
1969	24,177	42,015	42,015
1970	516	35,380	52,760
1971	0	130	10,598
1972	0	0	30
1973	0	0	0
Total Size	24,693	77,525	105,393

CHAPTER VI

AN AGE-STRATIFIED MODEL

Other Stratified Models

The model considered up to this point assumed homogeneous mixing among the population. However, this is not usually an accurate description of any population, especially a large one.

Abbey (10) recognized that

> if the population in which an epidemic occurs is actually a group of smaller populations with some cross-contact among them, this will introduce a source of variation not considered in the model.

Not having the necessary data available to test against a real population, she divided her population into two disjoint sub-populations A and B, with some mixing between them and calculated a theoretical epidemic based on the combined epidemics of A and B. Using a variation of the Reed-Frost model, Abbey experimented with selected within- and between-population contact rates.

The results showed little variation between the combined epidemics and the epidemic based on a single contact rate (the largest deviation produced a chi-square statistic at a significance level of .38). Abbey felt that in order to obtain more conclusive results,

> a more realistic model of the conditions likely to be found in an actual community is needed, for example, on giving each individual several contact rates.

Following this suggestion, Lila Elveback, et al. (19), divided a population into multiple mixing groups, each person belonging to three or four of these randomly mixing groups. These groups included the following: family, playgroup, cluster, grade school, and community. Each group has its own "level of intimacy of contact," each represented in the contact rate.

If the probability of infection is $1 - q_1^{C_1}$ where q_1 = the contact rate with group 1 and C_1 = the number of infectants in group 1, then the probability that a susceptible individual will escape infection during an interval is P_E, the product of terms such as $1 - q_1^{C_1}$ for each mixing group to which he belongs.

Looking only at a community of 100 people, Elveback's model considers each susceptible, one at a time by computing P_E for each. By selecting a random number r_i, the susceptibility of the individual is determined. If $r_i > P_E$ for susceptible i, the individual becomes infected; otherwise, the susceptibility remains unchanged and the next susceptible is considered.

Both Abbey (10) and Elveback, et al. (19), have dealt with non-homogeneous mixing groups using variations of the Reed-Frost model. The purpose of this chapter will be to present a variation of the standard deterministic equation

$$C_{t+1} = \beta_t C_t S_t \; / \; P_t \qquad (6.1)$$

in order to deal with disjoint groups in a community, specifically an age-stratified population.

The Model

Let the initial model population consist of the total number of persons in the population under study. However, it is now divided into seven age groups--less than 1 year, 1 to 4 years, 5 to 9 years, 10 to 14 years, 15 to 19 years, 20 to 39 years, and 40 and over. These divisions are justified by the fact that the immunity rates differ greatly between each of these groups (4).

Each of these age groups is again subdivided into three epidemiological classes, according to past history with rubella (based on age-stratified random serosurveys). The categories consist of infectants, susceptibles, and immunes. It will be assumed that age group 1 (less than 1 year old) is incapable of either transmitting or contracting rubella and therefore only groups 2 through 7 will be considered.

One approach to allocating infectants to each age group is to distribute them according to the usual age distribution of the disease. After computing C_t, the number of infectants at time t, let the number of infectants in age group j be represented by $C(j,t)$ so that

$$C(j,t) = \hat{p}_j \, C_t \qquad (j = 2, \ldots 7) \qquad (6.2)$$

where

$\hat{p}_j$ = an estimated proportion of the infectants found in age group j.

In this case $\hat{p}_j$ is a variable dependent upon prior knowledge about the distribution of rubella infection. Note that

$$\sum_j \hat{p}_j = 1.$$

A disqualifying characteristic of this approach is that it is possible to have new infectants in any group at time $t+1$ with no susceptibles in that group at time t.

Another method of dividing the infectant population into age groups is to allocate the infectants according to the proportion of the susceptible population in each group. In other words, again compute C_t and let $C(j,t)$ be represented by the following equation:

$$C(j,t) = \frac{S(j,t-1)}{S_{t-1}} C_t \qquad (j = 2,\ldots 7) \qquad (6.3)$$

where

$S(j,t-1)$ = the number of susceptibles in age group j at time t-1,

S_t = the total number of susceptibles at time t.

However, the distribution of new infectants from the susceptible population is dependent upon factors other than the distribution of susceptibles. As in the standard epidemic Equation 6.1

$$C_t = \beta S_{t-1} C_{t-1} / P_{t-1}$$

the contact rate between susceptibles and infectants, β, plays an important part in the epidemic process. Similarly, when considering age groups, that rate at which these groups interact have a significant effect on the spread of rubella.

Neither of these two methods allows for changes in environmental conditions that may alter the actual contact within a community. For instance, the closing of schools decreases the mixing of the largest susceptible and infectant group and therefore may affect the epidemic process. This occurrence cannot be incorporated into either Equation 6.2 or 6.3.

Consider the equation

$$C(j,t) = \beta_t S(j,t-1) \left(\sum_{i=1}^{7} a_{i,j} \frac{C(i,t-1)}{P(i,t-1)} \right) \qquad (6.4)$$

where

$P(i,t-1)$ is the number of persons in group i at time t-1, and

$a_{i,j}$ is the proportion of total contact that group j has with group i.

Since the population is not mixing homogeneously, the proportion of infectants with which an individual comes in contact is adjusted by the proportion of time the susceptible spends with each group.

Note that

$$\sum_{j=2}^{7} a_{i,j} = 1 \qquad \text{for } i = 2,\ldots 7 \qquad (6.5)$$

although it is <u>not</u> necessarily true that

$$\sum_{i=2}^{7} \alpha_{i,j} = 1 \qquad \text{for } j = 2,\ldots 7 \qquad (6.6)$$

since α represents the proportion of group j's time that is spent with group i. Also, the $\alpha_{i,j}$'s will be assumed constant over time unless otherwise specified.

Solving this age-dependent model is more complicated than the model described in Chapter V. The number of unknown variables has increased. There are now 36 $\alpha_{i,j}$'s ($i = 2,\ldots 7$ and $j = 2,\ldots 7$) needed before a solution can be found.

Before trying to find exact solutions to the $\alpha_{i,j}$'s, hypothetical values were chosen to test the simulation model of an age-stratified population. A listing of this program is shown in Appendix II. In this revised model, the population is divided into one-year age intervals for the purpose of aging as time is incremented. This distribution of the initial population is based on census data (22). The initial susceptible, immune and infectant populations are also divided into age intervals distributed according to previously discussed studies. In particular, the infectants were distributed into age groups according to an epidemiological study conducted by CDC (4) in which it was found that the p_j's (where $p_j = C_t / C(j,t)$) described in Equation 6.5 to be the following:

Age Group j	p_j	Cumulative p_j
1	0.000	0.000
2	.135	.135
3	.431	.566
4	.235	.801
5	.121	.922
6	.067	.989
7	.011	1.000

For the purpose of output and analysis, the one-year age intervals are accumulated into the seven age-groups. The $a_{i,j}$'s for these groups were chosen in an attempt to approach a steady state situation in which the distribution of infectants remained constant. The $a_{i,j}$'s are shown in Table 7.

TABLE 7

A Test Set of $a_{i,j}$'s

i \ j	2	3	4	5	6	7
2	8/72	4/72	4/72	4/72	16/72	36/72
3	8/72	28/72	16/72	8/72	8/72	4/72
4	6/72	14/72	19/72	17/72	10/72	6/72
5	4/72	4/72	14/72	22/72	22/72	6/72
6	6/72	8/72	12/72	14/72	18/72	14/72
7	2/72	2/72	2/72	2/72	2/72	60/72

The $a_{i,j}$'s listed in Table 7 and Equation 6.4 were incorporated into

the model as the method of determining the number of infectants from each group j for each time period t. Table 8 shows the results of simulation in terms of the distribution of infectants over time.

TABLE 8

Distribution of Infectants

Age Group	Time Period t		
	1	13	25
1	.0	.0	.0
2	.135	.312	.314
3	.431	.249	.246
4	.235	.181	.181
5	.121	.112	.112
6	.067	.070	.071
7	.011	.076	.076

It can be seen from this table that the initial steady state situation was not maintained. The equilibrium shifted almost immediately, however, and then was maintained. Note that deviations may actually occur in a real situation, but for the purpose of this model $a_{i,j}$'s that maintain the status quo are desired.

With this purpose in mind an attempt was made to solve for these 36 $a_{i,j}$'s ($i = 2,\ldots 7$ and $j = 2,\ldots 7$). To set up the initial problem, Equation 6.4 was expanded for all j so that there are six equations with 36 unknowns. Each equation contains six mutually exclusive $a_{i,j}$'s

creating six separate problems. The only other available knowledge about these contact rates is that listed in Equation 6.5, $\sum_j \alpha_{i,j} = 1$. Expanding Equation 6.4

$$C(j,t) = \alpha_{2,j} D_2 + \alpha_{3,j} D_3 + \alpha_{4,j} D_4 + \alpha_{5,j} D_5 + \alpha_{6,j} D_6 + \alpha_{7,j} D_7 , \quad (6.7)$$

where

$$D_i = \beta_t \, S(j, t-1) \, C(i, t-1) \, / \, D(i, t-1) .$$

In order to produce more equations containing these variables, Equation 6.4 was incremented through time. For 26 time periods the total number of infectants was calculated and assigned to age groups by substituting into Equation 6.1. For each time period t a value was calculated for C(i, t) for each i. Using simple algebraic techniques, S(i, t) and P(i, t) were computed using Equations 4.1 through 4.4.

There are now 27 simultaneous equations for each j (26 in the form of Equation 6.4 and one in the form of Equation 6.5) in which the $\alpha_{i,j}$'s are the variables. To solve these simultaneous equations, a multiple regression technique was used where C(j, t) is the dependent variable and the D_i's (i = 2, . . . 7) are the independent variables. However, when a multiple regression was applied, it was realized that the D_i's actually are not independent. This is a result of the method by

which they were generated. For any time t they were generated using the same deterministic equations. Furthermore, knowledge is enough information to solve for the remaining five D_i's.

The only way to compensate for this is to have actual data concerning the occurrence of real rubella cases for several time periods t. If actual data were available, the program listed in Appendix II gives a more realistic description of both a dynamic population and the rubella process occurring within that population than does the model described in Chapters IV and V.

Other studies have handled this problem by choosing an hypothetical population which could not be expected to mimic a pre-specified real world situation.

CHAPTER VII

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

In 1840, Farr wrote

> Epidemics have furnished much matter for discussion and still offer large scope for inquiry (7).

Since that time a mathematical approach to the theory of epidemics has developed. Using any of the empirical, deterministic, or stochastic techniques, such phenomena can be viewed as having a mathematical basis. In attempting to describe the relationship, there is a better understanding of the relationship of the factors affecting the epidemic process.

The regularity of rubella lends itself well to such mathematical description. The period of infection is both constant among all individuals and predictable, after which the infected individual recovers and becomes immune. The incidence of rubella varies widely from year to year with unusually high incidence every six to nine years.

Two deterministic, discrete-time-interval simulation models for rubella have been presented. Both models are based on Bailey's simple deterministic model which has been expanded by adding births, deaths, recovery from disease, and the possibility of vaccination. The

population under study is the East North Central section of the United States beginning in 1966.

The first model considers this population to be unstructured (there is no difference between individuals other than state of disease) and mixing within this total population is homogeneous. Basing the initial population distribution on the results of serosurveys and basing the contact rates on one year's reported rubella cases, the simulated results come reasonably close to predicting the actual reported cases. The largest discrepancy occurs in the latter part of 1968 predictions, the time at which vaccination programs were beginning to occur.

The problems involved in estimating the proportion of the population that is immune and the proportion of the population that is susceptible are many. For this reason the effect of variations in these proportions was tested. It was found that by increasing the initial proportion of immunes, a larger number of infectants were produced and the next major epidemic occurred at a sooner point in time. The opposite was true when the initial proportion of immunes was decreased.

Another difficult parameter to determine is the degree of under reporting of rubella cases. Therefore, a comparison was made by varying the under reporting correction factor. It was found that as the correction factor increases, the predicted number of cases also increases.

The advent of rubella vaccine has contributed another factor to

the epidemic process. Several alternative vaccination programs were compared. It was found that only 25 per cent of the susceptible population need be vaccinated in order to completely wipe out rubella infection from the population. Vaccinating any larger fraction produced diminishing returns. In addition to the extent of the vaccination program, timing also had an effect on the course of the epidemic.

The second model differs from the first in that the population is considered to be structured according to age groups, and mixing within this population is not homogeneous. The population was divided into seven age groups. Each group had a distinctly different proportion of immunes, different death rates and mixing between and within groups takes place with different contact rates. An original model describing this inter- and intra-group mixing was presented in Equation 6.4. Although the model was not solved so as to mimic actual occurrences, a test set of parameters was used to produce hypothetical results. It was found that a steady state situation, although not the initial distribution, was reached and maintained. This second model provides a more realistic description of a population than the first. However, in making the model more realistic, more information is needed concerning the population under study.

Recommendations

It is important to realize the limitations of mathematical

models. Although they are good exercises in understanding the epidemic process, it is sometimes difficult to fit them to a real population; the necessary data is often not available. Since most of the studies in this area are for hypothetical populations, more models should be based on real data from real populations. It is the recommendation of this author that, in order to continue research along the lines of this thesis, a smaller population should be chosen in which, if the necessary data were not available, it could feasibly be collected. The data should consist not only of epidemiological information, but also age-structure and population dynamics. It is such studies that indicate the inadequacies of the epidemiological and demographic information presently available.

APPENDICES

APPENDIX I

RUBELLA SIMULATION MODEL

```
      INTEGER POP,SUSC,SUMINF,BIRTHS,DEATHS
      DIMENSION PERYR(10),NVAC(10),NVACPD(10)
      DIMENSION IFACT(5),INREAL(1000),RT(250),FACTOR(53)
      DIMENSION POP(1000),SUSC(1000),INF(1000),IMMUNE(1000),
     1 PDFINF(1000),IDIFF(10),SUMINF(1000)
      DIMENSION NF1(250),INRL1(250),RLSICK(250),OBSSIC(250)
      DIMENSION RATE(210),NF(210),ITOTAL(20)
      READ(5,10) (IFACT(JJ),JJ=1,5)
   10 FORMAT(5I5)
      IX=1
      IX1=IX+1
      IW=26
      IX52=IW*IX
      IX521=IX52+1
      ITIMES=1
      READ(5,19) (NF(JJ),JJ=1,53)
      READ(5,19) (NF(JJ),JJ=53,105)
      READ(5,19) (NF(JJ),JJ=105,157)
      READ(5,19) (NF(JJ),JJ=157,209)
   19 FORMAT(16I5)
      CALL GROUP(NF,NF1,208,2,IW)
      DO 1001 IT=1,ITIMES
C
C THIS ROUTINE READS IN DATA
C
      READ(5,1) POP(1),IMMUNE(1),BIRTH,DEATH,BETA1
    1 FORMAT(2I10,20X,3F10.0)
      BIRTH=BIRTH/1000.
      W=IW
      BIRTH=EXP(ALOG(1.+BIRTH)/ W )-1.
      DEATH=DEATH/1000.
      DEATH=1.-EXP(ALOG(1.-DEATH)/W)
      DO 1002 JJJ=1,1
      IX521=IX52+1
      DO 100 JJ=1,IX521
      INF(JJ)=IFACT(JJJ)*NF1(JJ)
  100 CONTINUE
      SUSC(1)=POP(1)-INF(1)-IMMUNE(1)
  150 BETA=BETA1
      DO 974 IVACIN=1,4
      READ(5,9)PCTVAC,(PERYR(IVYR),IVYR=4,10)
    9 FORMAT(8F3.2)
      DO 12 IVYR=1,3
      NVACPD(IVYR)=0
   12 NVAC(IVYR)=0
```

```
      IYEAR=1
      SUMINF(1)=INF(1)
      NYRS=7
      IWW=NYRS*IW
      DO 1000 J=1,IWW
      N=POP(J)
      XN=N
      BIRTHS=BIRTH*XN
      DEATHS=DEATH*XN
      IF(IYEAR.GE.IX1) GO TO 450
      IF(INF(J).EQ.0) GO TO 1000
      IF(SUSC(J).EQ.0) GO TO 1000
      IF(J.EQ.1) GO TO 420
      INF(J)=FLOAT(INF(J))*(1.-DEATH)
      SUSC(J)=FLOAT(SUSC(J))*(1.-DEATH)
      IMMUNE(J)=FLOAT(IMMUNE(J))*(1.-DEATH)
       POP(J)=POP(J)-DEATHS
  420 RATE(J)=FLOAT(INF(J+1))/FLOAT(INF(J))*FLOAT(POP(J))/
     1FLOAT(SUSC(J))
      GO TO 500
C
C THIS ROUTINE FINDS THE STATUS OF TIME PERIOD J + 1
C
  450 IWK=MOD(J,IX52)
      IF(IWK.EQ.0) IWK=IX52
      IF(IYEAR.NE.4.OR.IWK.NE.1) GO TO 460
      TOTVAC=PCTVAC*FLOAT(SUSC(J))
      DO 455 IVYR=4,NYRS
      NVAC(IVYR)=PERYR(IVYR)*FLOAT(SUSC(J))
      NVACPD(IVYR)=PERYR(IVYR)*FLOAT(SUSC(J))/W
  455 CONTINUE
  460 CONTINUE
      SUSC(J)=FLOAT(SUSC(J))*(1.-DEATH)
      INF(J)=FLOAT(INF(J))*(1.-DEATH)
      IMMUNE(J)=FLOAT(IMMUNE(J))*(1.-DEATH)
       POP(J)=POP(J)-DEATHS
      IF(SUSC(J).LT.NVACPD(IYEAR)) NVACPD(IYEAR)=0
      SUSC(J)=SUSC(J)-NVACPD(IYEAR)
      INF(J+1)=RATE(IWK)*FLOAT(INF(J))/FLOAT(POP(J))*
     1FLOAT(SUSC(J))+.5
      IF(INF(J+1).GE.POP(J)) GO TO 1001
  500 SUSC(J+1)=SUSC(J)-INF(J+1)+BIRTHS
      IMMUNE(J+1)=IMMUNE(J)+INF(J)+NVACPD(IYEAR)
      SUMINF(J+1)=SUMINF(J)+INF(J+1)
      POP(J+1)=POP(J)+BIRTHS
C
C THIS ROUTINE WRITES OUT THE STATUS OF TIME PERIOD J + 1
C
  501 JMOD52=MOD(J,IW)
      IF(JMOD52.EQ.0) JMOD52=IW
      INREAL(J)=INF(J)/IFACT(JJJ)
```

```
      IF(MOD(J,IW).NE.0) GO TO 1000
      IF(J.EQ.IW) IDIFF(1)=SUMINF(IW)
      ITEMP=J-IW
      IF(J.NE.IW) IDIFF(IYEAR)=SUMINF(J)-SUMINF(ITEMP)
      IYEAR=J/IW+1
 1000 CONTINUE
      WRITE(6,24)
   24 FORMAT(1H1)
      WRITE(6,20) (L,L=1,NYRS)
   20 FORMAT(////,7X,5HRATE , 7(3X,4HYEAR,I3),/)
      DO 800 LL=1,IW
      WRITE(6,21) RATE(LL),(INREAL(L),L=LL,IWW,IW)
   21 FORMAT(2X,F10.3,10I10,F10.3)
  800 CONTINUE
      DO 900 LL=1,NYRS
      ITOTAL(LL)=IDIFF(LL)/IFACT(JJJ)
  900 CONTINUE
      WRITE(6,22) (ITOTAL(L),L=1,NYRS)
   22 FORMAT(3X,9HT O T A L,10I10)
      WRITE(6,5) BIRTH,DEATH,IFACT(JJJ)
    5 FORMAT(/,10X,11HBIRTH RATE=,F9.6,5X,11HDEATH RATE=,
     1F9.6,5X,7HFACTOR=,I5)
      WRITE(6,29)(IMMUNE(L),L=1,IWW,IW)
   29 FORMAT(2X,7HIMMUNES,3X,10I10)
      WRITE(6,23) (POP(L),L=1,IWW,IW)
   23 FORMAT(2X,10HPOPULATION,10I10)
      WRITE(6,24)
      WRITE(6,31) (NVAC(IVYR),IVYR=1,NYRS)
   31 FORMAT(2X,10HVACCINATED,10I10)
  974 CONTINUE
  975 CONTINUE
 1002 CONTINUE
 1001 CONTINUE
      CALL EXIT
      END
       SUBROUTINE GROUP (IX,IXM,L,M,IW)
       DIMENSION IXM(250),IX(1000)
C
C   THIS ROUTINE SUBTOTALS ARRAY INTO FOUR WEEK PERIODS
C
       N=1
       DO 100 J=1,L,M
       IXM(N)=0
       DO 50 I=1,M
       ITEMP=J+I-1
       IXM(N)=IXM(N)+IX(ITEMP)
    50 CONTINUE
       N=N+1
   100 CONTINUE
       RETURN
       END
```

APPENDIX II

AGE-STRATIFIED RUBELLA SIMULATION MODEL

```
      INTEGER YEARS,FACTOR,GROUP,YOUNG,OLD
      INTEGER POP,SUSC,SUMINF,BIRTHS,DEATHS
      REAL MORTAL, IMMRT,NOTWEL,ILL,MXRATE,MIX,INFECT,IWELL
      DIMENSION ILL(86),MXRATE(7,7),RATE(210),X(86),NF(54)
      DIMENSION AREA(6,18),SICGRP(7),FACTOR(53)
      DIMENSION AGEGRP(16),GROUP(7)
      DIMENSION YEARS(7),MORTAL(90),SICKRT(86),IMMRT(86),
     1         AGEPCT(86),INFECT(86),IWELL(86),SUSCEP(86),
     1         SUSGRP(7),WELGRP(7),PEOPLE(86),INFGRP(7)
      DIMENSION IMMGRP(520),ISUGRP(520),POP(520),SUSC(520),
     1IMMUNE(520),SUMINF(520),IDIFF(10),INF(520)
      DATA AGEGRP/4H  0-,2H 1,4H  1-,2H 4,4H  5-,2H 9,
     1             4H 10-,2H14,
     1             4H 15-,2H19,4H 20-,2H39,4H 40 ,2H+ ,
     1             4H TOT,2HAL/
      REWIND 8
      DEATHS=0
      BIRTHS=0
      LASTWK=54
      FACTOR(1)=20
      READ(5,17) (AGEPCT(L),L=1,86)
   17 FORMAT(25F3.2)
C
C THIS ROUTINE SETS UP TRANSITION RATES FOR EACH AGE GROUP
C              AND AGE LEVEL
C
      IYR=1
      READ(5,18) (MORTAL(L),L=1,5)
      READ(5,18) (MORTAL(L),L=6,86,5)
   18 FORMAT(16F5.0)
      DO 50 LTIME1 =6,86,5
      DO 50 LTIME2 =1,4
      L=LTIME1 +LTIME2
      MORTAL(L)=MORTAL(LTIME1)
   50 CONTINUE
      DO 60 L=1,86
      MORTAL(L)=MORTAL(L)/1000.
      MORTAL(L)=EXP(ALOG(1.+MORTAL(L))/52.)-1.
   60 CONTINUE
      WRITE(6, 2)
    2 FORMAT(//,2X,8HAGE SPAN,2X,14HINFECTION RATE,2X,
     1   13HIMMUNITY RATE,/)
C
```

```
C
C THIS ROUTINE FINDS DEATHS AT EACH AGE LEVEL
C
  301 DEATHS=0
      DO 400 L=1,86
      INFECT(L)=(1.-MORTAL(L))*      INFECT(L)
      SUSCEP(L)=(1.-MORTAL(L))*SUSCEP(L)
      IWELL(L)=(1.-MORTAL(L))*IWELL(L)
      DEATHS=MORTAL(L)*(INFECT(L)+SUSCEP(L)+IWELL(L))+DEATHS
  400 CONTINUE
C
C THIS ROUTINE AGES POPULATION
C
      INFECT(L)=(1.-1./52.)*INFECT(L)+(1./52.)*INFECT(L-1)
      SUSCEP(1)=(1.-1./52.)*SUSCEP(1)+FLOAT(BIRTHS)
      INFECT(1)=(1.-1./52.)*INFECT(1)
      IWELL(1) =(1.-1./52.)*IWELL(1)
      PEOPLE(1)=INFECT(1)+SUSCEP(1)+IWELL(1)
      SICK=0.
      DO 500 L=2,85
      BIRTHS=BIRTH*FLOAT(POP(J-1))
      SUSCEP(L)=(1.-1./52.)*SUSCEP(L)+(1./52.)*SUSCEP(L-1)
      IWELL (L)=(1.-1./52.)*IWELL (L)+(1./52.)*IWELL (L-1)
      PEOPLE(L)=INFECT(L)+SUSCEP(L)+IWELL(L)
      SICK=SICK+SUSCEP(L)
  500 CONTINUE
      INFECT(86)=INFECT(86)+(1./52.)*INFECT(85)
      SUSCEP(86)=SUSCEP(86)+(1./52.)*SUSCEP(85)
      IWELL (86)=IWELL (86)+(1./52.)*IWELL (85)
      PEOPLE(86)=INFECT(86)+SUSCEP(86)+IWELL(86)
      SICK=SICK+SUSCEP(86)
C
C   THIS ROUTINE FINDS INFECTION RATE FOR EACH MONTH
C    OF KNOWN YEARS (SEASONAL VARIATION)
C
      IF(J.GT.53) GO TO 570
      RATE(J-1)=FLOAT(INF(J))/FLOAT(INF(J-1))*
     1  FLOAT(POP(J-1))/FLOAT(SUSC(J-1))
      BETA=RATE(J-1)
      GO TO 580
  570 JMOD52=MOD(J,52)
      IF(JMOD52.EQ.0) JMOD52=52
      BETA=RATE(JMOD52-1)
C
C THIS ROUTINE INFECTS, RECOVERS, ETC. POPULATION
C
      INF(J)=(BETA*FLOAT(INF(J-1))/FLOAT(POP(J-1)))*
     1 FLOAT(SUSC(J-1))+.5
  580 NOTWEL=0.
      IYR1=2
      L=2
```

```
      DO 585 LL=2,7
      ILL(LL)=SICKRT(L)*FLOAT(INF(J))*YEARS(LL)
      L=L+YEARS(LL)
  585 CONTINUE
      L=2
      DO 600 LL=2,7
      ISPAN=YEARS(LL)
      DO 600 LLL=1,ISPAN
      INFECT(L)=ILL(LL)*SUSCEP(L)/SUSGRP(LL)
      IWELL(L)=IWELL(L)+INFECT(L)
      SUSCEP(L)=SUSCEP(L)-INFECT(L)
      L=L+1
  600 CONTINUE
C
C  THIS ROUTINE PLACES PEOPLE IN PROPER AGE GROUPS
C    AND ILLNESS CATEGORIES
C
  601 POP(J)=0
      IYR=1
      IMMUNE(J)=0
      SUSC(J)=0
      DO 700 LL=1,7
      SICGRP(LL)=0.
      SUSGRP(LL)=0
      WELGRP(LL)=0
      ISPAN=YEARS(LL)
      DO 650 LLL=1,ISPAN
      SICGRP(LL)=SICGRP(LL)+INFECT(IYR)
      SUSGRP(LL)=SUSGRP(LL)+SUSCEP(IYR)
      WELGRP(LL)=WELGRP(LL)+IWELL(IYR)
      IYR=IYR+1
  650 CONTINUE
      ISUGRP(LL)=SUSGRP(LL) + .5
      IMMGRP(LL)=WELGRP(LL) + .5
      INFGRP(LL)=SICGRP(LL) +.5
      GROUP(LL)=IMMGRP(LL)+ISUGRP(LL)+INFGRP(LL)
      POP(J)=POP(J)+GROUP(LL)
      IMMUNE(J)=IMMUNE(J)+IMMGRP(LL)
      SUSC(J)=SUSC(J)+ISUGRP(LL)
  700 CONTINUE
C
C  THIS ROUTINE WRITES OUT POPULATION STATUS FOR
C    EACH TIME PERIOD
C
  800 WRITE(6,13)
   13 FORMAT(/,2X,8HAGE SPAN,2X,
     1              11H POPULATION,2X,10HINFECTANTS,2X,
     1  12HSUSCEPTIBLES,
     23X,7HIMMUNES,8X,4HSICK,3X,9HSICK RATE,1X,
     2  11HIMMUNE RATE,/)
```

```
C  THIS ROUTINE READS AND WRITES INITIAL INFECTION RATES
C       AND IMMUNITY RATES
C
      DO 100 LL=1,7
      READ(5,9) YOUNG, OLD, SICK,WELL
    9 FORMAT(2I5,2F5.0)
      YEARS(LL)=OLD-YOUNG
      ISPAN=YEARS(LL)
      DO 80 LLL=1,ISPAN
      SICKRT(IYR)=SICK/ISPAN
      IMMRT(IYR)=WELL
      IYR=IYR+1
   80 CONTINUE
      MM1=2*LL-1
      MM2=2*LL
      WRITE(6,12) AGEGRP(MM1),AGEGRP(MM2),SICK,WELL
   12 FORMAT(5X,A4,A2,6X,F10.3,5X,F10.3)
  100 CONTINUE
      J=1
C
C  THIS ROUTINE READS IN INITIAL POPULATION DATA
C
      READ(5,3) NPLACE,(AREA(  1,M),M=1,18)
    3 FORMAT(I5,2X,18A4)
      READ(5,1) POP(1),IMMUNE(1),INF(1),SUSC(1),BIRTH,DEATH,
     1  BETA1
    1 FORMAT(4I10,3F10.0,F2.0,F8.0)
      READ(5,19) (NF(JJ),JJ=1,54)
   19 FORMAT(16I5)
      DO 150  JJ=1,54
      INF(JJ)=FACTOR(1) *NF(JJ)
  150 CONTINUE
      SUSC(J)=POP(J)-IMMUNE(J)-INF(J)
      RATE(1)=1.
      BETA=RATE(1)
      WRITE(6,5) (AREA(1,M),M=1,18)
    5 FORMAT(1H1,10X,18A4,//)
      IYEAR=1
      WRITE(6,4)IYEAR
    4 FORMAT(1H0,20X,8HY E A R ,I3,//)
      SUMINF(1)=INF(1)
C
C THIS ROUTINE SETS INITIAL STATUS FOR EACH AGE GROUP
C
      DO 200 L=1,86
      PEOPLE(L)=AGEPCT(L)*FLOAT(POP(1))/100.
      INFECT(L)=SICKRT(L)*FLOAT(INF(1))
      IWELL(L)=IMMRT(L)*PEOPLE(L)
      SUSCEP(L)=PEOPLE(L)-IWELL(L)-INFECT(L)
  200 CONTINUE
  201 GO TO 601
```

```
      DO 810 LL=1,7
      REALL=FLOAT(INFGRP(LL))/FACTOR(1)
      SIC=FLOAT(INFGRP(LL))/FLOAT(INF(J))
      WEL=FLOAT(IMMGRP(LL))/FLOAT(GROUP(LL))
      MM1=2*LL-1
      MM2=2*LL
      WRITE(6,14) AGEGRP(MM1),AGEGRP(MM2),GROUP(LL),
     1  INFGRP(LL),
     1ISUGRP(LL),IMMGRP(LL),REALL,SIC,WEL
   14 FORMAT(4X,A4,A2,  3X,I10,2X,I10,4X,2I10,2X,F10.2,2X,
     1  F10.3,2X,F10.3)
  810 CONTINUE
      WRITE(6,15)AGEGRP(15),AGEGRP(16),POP(J),INF(J),
     1   SUSC(J),IMMUNE(J)
   15 FORMAT(4X,A4,A2,3X,I10,2X,I10,4X,2I10)
      WRITE(8) (ISUGRP(LL1),LL1=2,7),(INFGRP(LL2),LL2=2,7),
     1  (GROUP(LL3), LL3=2,7),BETA
      JMIN1=J-1
      IF(JMIN1.LE.1) JMIN1=1
      WRITE(6,16) BIRTHS,DEATHS,BETA,J
   16 FORMAT(/,2X,7HBIRTHS=,I10,2X,7HDEATHS=,I10,20X,F8.3,
     1  I6,/)
 1000 J=J+1
      IF(J.EQ.LASTWK) CALL EXIT
      IF(MOD(J,52).EQ.1) IYEAR=IYEAR+1
      IF(MOD(J,52).EQ.1) WRITE(6,4) IYEAR
      GO TO 301
 1001 REWIND 8
      CALL EXIT
      END
```

LITERATURE CITED

(1) L. Z. Cooper and S. Krugman, The Rubella Problem, Yearbook Medical Publishers, Inc., 1969.

(2) B. D. Davis, et al., Microbiology, Harper and Row Publishers, 1969, pp. 1,368-1,372.

(3) J. J. Witt, et al., "Epidemiology of Rubella," American Journal Dis. Child., Vol. 118, July, 1969, p. 107

(4) Rubella Surveillance, National Communicable Disease Center publication, No. 1, June, 1969.

(5) "First Rubella Vaccine OK'd," Journal of the American Medical Association, Vol. 208, June 16, 1969, p. 2,004.

(6) "Shortage of Vaccine Funds May Slow Rubella Campaign," JAMA, Vol. 210, October 13, 1969, p. 233.

(7) R. E. Serfling, "Historical Review of Epidemic Theory," Human Biology, Vol. 24, September, 1952, pp. 145-166.

(8) N. T. J. Bailey, The Mathematical Theory of Epidemics, Griffin & Company Ltd., London, 1957.

(9) W. O. Kermack and A. G. McKendrick, "A Contribution to the Mathematical Theory of Epidemics," Proceedings of the Royal Society of London, Vol. 115A, 1927, p. 700.

(10) H. Abbey, "Examination of Reed-Frost Theory," Human Biology, Vol. 24, September, 1952, pp. 201-233.

(11) R. E. Serfling, Mathematical Models in Epidemiology and Demography, Unpublished Class Notes, Tulane School of Public Health and Tropical Medicine, Fall, 1971.

(12) M. S. Bartlett, "Deterministic and Stochastic Models for Recurrent Epidemics," Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. 4, 1956, pp. 81-109.

(13) N. T. J. Bailey, "A Simple Stochastic Epidemic," Biometrika, Vol. 37, 1950, pp. 193-202.

(14) M. E. Thomas, "A Stochastic Model for the Analysis of Epidemics," Proceedings of the American Astronautical Society, June, 1969.

(15) C. Revelle, F. Feldman, and W. Lynn, "An Optimization Model of Tuberculosis Epidemiology," Management Science, Vol. 16, No. 4, December, 1969, pp. 190-211.

(16) H. Waaler, A. Geser, and S. Anderson, "The Use of Mathematical Models in the Study of the Epidemiology of Tuberculosis," American Journal of Public Health, Vol. 52, No. 6, June, 1962, p. 1,002.

(17) S. Ferebee, "Is Eradication Too Much Trouble?" National Tuberculosis Association Meeting, May, 1966.

(18) "Rubella Virus Vaccine," Morbidity and Mortality Weekly Report, Vol. 20, No. 34, August 28, 1971, pp. 304-305.

(19) L. Elveback, E. Ackerman, L. Gatewood, and J. P. Fox, "Stochastic Two-Agent Epidemic Simulation Models for a Community of Families," American Journal of Epidemiology, Vol. 93, No. 4, April, 1971, pp. 267-280.

(20) Vital Statistics of the United States, 1967, U. S. Department of Health, Education, and Welfare, Washington, 1969.

(21) R. Wallace, National Communicable Disease Center, Immunization Branch, Personal Communication, 1970.

(22) Current Population Reports, U. S. Department of Commerce, Bureau of the Census, Series P-25, December, 1967.

OTHER REFERENCES

Abrutyn, E., National Center for Disease Control, Investigation and Evaluations Section, Personal Communication, 1969-1970.

Brogger, S., "Systems Analysis in Tuberculosis Control: A Model," American Review of Respiratory Diseases, Vol. 95, 1967, pp. 419-434.

Elveback, L., et al., "An Extension of the Reed-Frost Epidemic Model for the Study of Competition between Viral Agents in the Presence of Interference," American Journal Hyg., Vol. 80, 1964, pp. 356-364.

Elveback, L., et al., "A Stochastic Model for Competition between Viral Agents in the Presence of Interference. I. Live Virus Vaccine in a Randomly Mixing Population, Model III," Amer. J. Epidemiology, Vol. 87, 1968, pp. 373-384.

"First Rubella Vaccine License May Be Issued Next Month," JAMA, Vol. 208, April 21, 1969, p. 454.

Goffman, "Stability of Epidemic Processes," Nature, Vol. 210, May, 1966.

Halstead, S. B., Char, D. F., and Diwan, A. R., "Evaluation of Three Rubella Vaccines in Adult Women," JAMA, Vol. 211, February 9, 1970, pp. 991-993.

Ingalls, T. H., "German Measles (1900-1960)," Archives of Environmental Health, Vol. 5, December, 1962, pp. 60-66.

Kendall, D. G., "Deterministic and Stochastic Epidemics in Closed Populations," Proc. Third Berkeley Symposium on Math. Statistics and Prob., Vol. 4, 1956, p. 149.

"Maternal Rubella," JAMA, Vol. 208, May 26, 1969, p. 1,484.

Neyman, J., and Scott, E., "Stochastic Models of Epidemics," Stochastic Models in Medicine and Biology, Madison, 1964.

"Proceedings of the International Conference on Rubella Immunization 1969," Amer. J. Dis. Child., July, August, 1969.